

# 척추 질환의 진단 및 결과 예측 분야에 대한 인공지능 학습법의 적용

최성훈 · 윤종현\* · 박예수<sup>†‡</sup>

한양대학교 의과대학 한양대학교병원 정형외과학교실, \*한양대학교 공과대학 ERICA 기계공학부, †한양대학교 의과대학 한양대학교구리병원 정형외과학교실

## A Narrative Review on the Application of Artificial Intelligence on the Diagnosis and Outcome Prediction for Spinal Diseases

Sung Hoon Choi, M.D., Ph.D., Jonghun Yoon\*, Ph.D., and Ye-Soo Park, M.D., Ph.D.<sup>†‡</sup>

Department of Orthopedic Surgery, Hanyang University Seoul Hospital, Hanyang University College of Medicine, Seoul,

\*Department of Mechanical Engineering, Hanyang University, ERICA, Ansan, †Department of Orthopedic Surgery, Hanyang University Guri Hospital, Hanyang University College of Medicine, Guri, Korea

Artificial intelligence is a concept that includes machine learning and artificial neural networks and is characterized by the rapid analysis of large-scale data based on fast computation of computers. Various medical studies have been reported to estimate the diagnosis and prognosis of specific diseases using artificial intelligence technology. Recently, many medical studies using a convolutional neural network, which is particularly useful for imaging recognition and analysis, have been reported. For various diseases in orthopedic surgery, including spinal diseases, imaging data, such as simple radiographs, computed tomography, and magnetic resonance imaging, are essential for determining the diagnosis and treatment of patients. Therefore, compared to other fields of medicine, there are many disease entities where artificial intelligence can be applied to the diagnosis and treatment of orthopedic diseases. With the increasing importance of medical informatics, the introduction of artificial intelligence by orthopedic surgeons will positively contribute to the development of medicine. This review outlines artificial intelligence and introduces studies using machine learning to diagnose and predict spinal diseases.

**Key words:** machine learning, artificial neural network, convolutional neural network, spine, diagnosis

### 서론

인간은 양 상지를 자유롭게 사용할 수 있는 완전한 직립 보행을 하며 요추의 전만을 보이는 유일한 종이다. 두개골을 떠받치고 무게 중심을 하지에 전달하는 척추는 전방에서는 수직이며, 측면에서는 연속적인 만곡을 보이는 특징을 보인다. 인간의 고유한

척추 시상면 정렬 상태는 균형대(cone of economy) 안에서 최소한의 에너지를 사용하여 기립 상태를 유지할 수 있도록 하게 한다. 척추 시상면 정렬을 이해하는 데 가장 중요한 영상학적 지표는 골반 입사각(pelvic incidence)이며 형태학적 지표(morphologic parameter)로서 각 개인은 일정 범위 안에서 고유한 값을 갖는다. 골반 입사각은 골반의 크기를 대변하는 지표이며 골반의 기울기인 천추 경사각(sacral slope)과 골반 경사각(pelvic tilt)의 합과 같다. 인간의 척추의 정렬 상태는 추간판과 후관절의 퇴행성 변화, 근감소증 및 노화에 의하여 요추의 전만각이 감소하면서 균형대에서 점차 멀어지는 변화를 겪게 되며, 골반의 후방 회전, 흉추의 후만 감소, 경추의 과전만, 슬관절의 굴곡 및 족관절의 신전 등의 보상 기전이 발생함이 알려져 있다.<sup>1,2)</sup> 이를

Received July 12, 2021 Revised September 5, 2021

Accepted October 13, 2021

Correspondence to: Ye-Soo Park, M.D., Ph.D.

Department of Orthopedic Surgery, Hanyang University Guri Hospital, 153 Gyeongchun-ro, Guri 11923, Korea

TEL: +82-31-560-2187 FAX: +82-31-557-8781 E-mail: hyparkys@hanyang.ac.kr

ORCID: <https://orcid.org/0000-0002-2801-2674>

교정하기 위한 성인 척추 변형(adult sagittal deformity) 환자의 변형 교정 수술에서 수술 범위의 결정, 예후 및 인접 분절 질환(adjacent segment disease) 등의 합병증을 예측하는 것은 많은 척추 의사들의 주된 관심사 중 하나이다.<sup>3,4)</sup>

인공 신경망(artificial neural network)은 인공 지능(artificial intelligence)의 한 종류로 기계 학습(machine learning)의 하위 분류에 속하며, 인간의 신경계를 본떠 생의학 분야의 다양한 비선형적인 문제를 해결하려는 수학적 접근 방식의 분석이다(Fig. 1). 인공신경망은 지정된 입력변수와 출력 변수 간의 관계를 학습하는 데 능숙하도록 고안된 기계 학습을 바탕으로 주로 예측 및 분류와 관련된 패턴 인식을 하는 분야에 광범위하게 적용되고 있다.<sup>5)</sup> 인간은 스스로 판단할 수 있는 인지 기능을 바탕으로 새로운 자극에 대하여 스스로 판단하여 의사를 결정한다. 인공 신경망 역시 인간의 학습 과정과 유사하며 컴퓨터 성능의 발전을 토대로 군집화(clustering), 구조화(hierarchy) 등의 방법을 통하여 데이터를 분류하고, 샘플 관찰을 통해 인간이 배우는 것과 같은 방식으로 학습한다. 초기 학습을 위한 데이터셋의 결과를 알고 있는 상태에서 알고리즘을 형성하고, 새로운 데이터를 입력하였을 때 예상되는 결과를 도출하는 기계학습법은 지도 학습(supervised learning)에 해당하며, 우주 천체의 정보 등 샘플 데이터의 결과를 알지 못하는 상태에서 초기 학습부터 정해진 결과값이 없이 인공신경망이 스스로 학습하는 경우 비지도 학습(unsupervised learning)에 해당한다. 일부 유전체 연구를 제외한 대부분의 의학연구의 인공신경망은 지도학습에 해당하며 비지도학습의 대표적인 예로는 군집화 알고리즘(clustering algorithm)을 사용하여 개인에게 맞춤형 뉴스를 제공하는 인터넷 포털의 인공지능이 해당한다.

기계 학습은 알고리즘을 설계하는 프로그래머에 의하여 설계된 함수를 반복 조건(for loop)을 활용하여 대규모의 데이터를 단기간에 처리하여 결과를 도출하는 프로그램이다. 기계학습의 하위 개념인 인공신경망 역시 함수로 미리 정의된 계산 규칙에 따라 서로 상호 작용하여 학습하며 이 각 함수의 연결이 인공 뉴런 그룹이다. 정해진 함수 알고리즘에 따라 입력 데이터와 출력 데

이터의 쌍으로 구성된 샘플 데이터를 전달하면 새로 가설을 세울 입력-출력 관계 패턴을 추정할 수 있는 방법으로 구조를 수정하며 자가 학습한다. 이 학습 과정이 끝나 생성된 알고리즘은 새롭게 입력된 데이터에 대한 결과를 예측하는 가설을 세우거나 예후를 예측하는 데 사용될 수 있다.<sup>6)</sup> 의학 정보 중 영상 정보는 미세한 픽셀의 수정으로 제공 가능한 데이터 셋이 무한하여 인공지능을 활용한 진단 및 분류가 용이하다. 따라서 많은 전문가들은 일반적으로 인공 신경망이 의학 연구 방법을 크게 변화시킬 수 있는 혁신적인 기술이라는 데 동의한다.

이러한 인공신경망은 컴퓨터 성능의 향상에 따라 본질적으로 빠르고 스스로 진화할 수 있는 특징이 있어 영상 지표를 수기 측정하여 척추 시상면 정렬을 분석하는 것에 비하여 빠르고 대규모 정보를 처리할 수 있는 장점이 있다. 최근에는 높은 정확성과 일치성을 보이는 인공신경망 연구들이 보고되면서 영상 지표의 수동 측정 등으로 인한 임상주의 시간 소모를 줄여줄 수 있는 도구로 각광을 받고 있다. 본 연구에서는 척추 질환에 대한 기계 학습의 발전, 임상 적용 및 척추 정렬에 대한 연구에 대해 고찰한다. 먼저 인공 신경망에 대한 간략한 일반 개요를 제시하고, 척추 연구에 실용적이거나 잠재적인 영향을 미치는 인공 신경망과 최근 발전에 대하여 기술한다.

## 본 론

### 1. 기계 학습 및 인공 신경망

가장 빠르게 발전하고 있는 인공 신경망 분야는 딥 러닝(deep learning)과 컨볼루션 신경망(convolutional neural network)이다. 대부분의 인공 신경망의 학습 방법은 심층 신경망, 즉 여러 계층의 네트워크 구조를 기반으로 한다. 이러한 인공 신경망은 이미지 처리, 음성 인식 및 자연어 처리와 같은 연구 분야에 큰 발전을 가져오고 있다. 최근 비약적인 발전을 보이고 있는 인공 신경망은 향상된 컴퓨터 성능뿐만 아니라 이미징(imaging) 데이터베이스 등을 비롯한 다양한 원천 자료에서 수집된 방대한 데이터셋의 활용 가용성에 영향을 받는다.

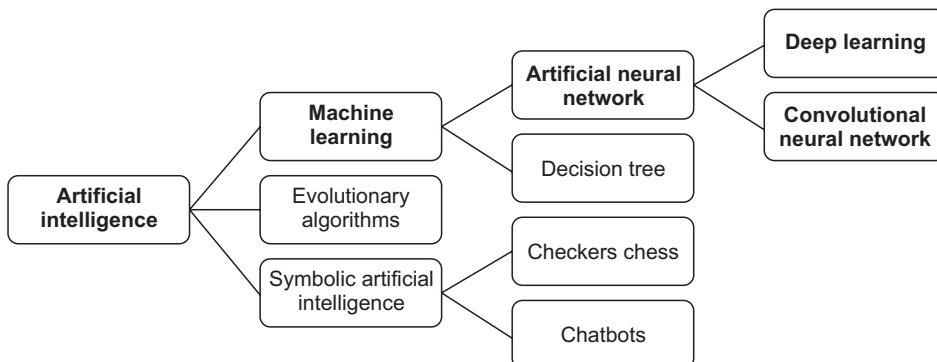


Figure 1. Schematic diagram of artificial intelligence on the research of spinal disorders.

“기계 학습”이라는 표현은 1959년 Samuel에 의해 도입되었으며, 그는 이를 명시적으로 프로그래밍하지 않고도 컴퓨터가 스스로 학습할 수 있는 능력을 부여하는 연구 분야로 정의하였다.<sup>7)</sup> 기계 학습의 일반적인 목표는 특정 주제에 대한 예측과 추정이다. 다양한 형태의 기계 학습 중 대부분의 의학 연구에서 사용되는 학습 방법은 지도학습으로, 수행할 작업의 특성에 따라 프로그램은 정확한 출력(실측 값)이 알려진 입력 모음을 기반으로 하여 출력을 예측하는 방법이다.<sup>8)</sup> 대부분의 기계 학습의 지도 학습 방법은 예측과 실측 값 간의 차이를 나타내는 손실 함수의 값을 최소화하여 입력을 출력에 연결하는 최적의 방법을 학습하는 것으로 구성된다.

인공 신경망은 최근 몇 년 동안 가장 인상적인 개선을 보인 기계 학습의 한 분야로 대개 두 용어는 동의어로 쓰이고 있으며, 척추 분야를 포함한 많은 의학 분야에서 인공 신경망을 적용한 다양한 연구들이 보고되어 있다. 인공 신경망은 실제 생체의 뉴런이 뇌에서 연결되고 상호 작용하는 방식을 모방한 학습법으로 생체에서 영감을 받은 네트워크 학습법이다. Hebbian 학습의 원리를 모방하여 정보는 각 층으로 구성되고 간단한 작업을 수행하는 인공 뉴런을 통해 입력층에서 출력층으로 전환된다.<sup>9)</sup> 입력된 데이터에 의하여 반복 구문을 실행하는 알고리즘의 집합체인 네트워크에 오류를 최소화하기 위하여 입력 값에 가중치를 곱한 선형 조합을 만들고 선형 또는 비선형 활성화 함수를 통해 결과를 처리한다(Fig. 2). 기계학습은 초기에 주어진 데이터가 풍부하다면 신뢰할 만한 결과가 도출되나, 주어진 데이터가 충분하지 않은 경우에도 도출된 결과는 해당 데이터에서는 잘 맞으나, 새로운 데이터로 분석하였을 때 잘못된 결과를 도출하는 오류를 범하게 된다. 따라서 이러한 오류를 줄이기 위하여 초기 입력 값에 서로 다른 가중치를 매겨 최적의 예측을 할 수 있도록 설계한다. 또한 특정 데이터셋에 알고리즘이 결정되는 것을 방지하기 위하여

네트워크에는 페널티 계수를 통해 가중치의 큰 값에 페널티를 적용하여 과적합의 위험성을 줄이고자 한다. 이러한 패턴의 반복 훈련을 통하여 인공 신경망의 훈련은 가중치의 최적 값을 구하는 것으로, 훈련 데이터에 속하는 입력이 처리되고 계층을 통해 전송되어 최적의 결과를 도출하는 출력이 생성된다.

인공신경망은 다른 뉴런 배열과 학습 방법을 기준으로 다른 유형으로 구분된다. 가장 일반적인 유형은 다층 퍼셉트론(multi-layer perceptron, MLP)이라고 하며 총 3개 이상의 층으로 구성되며, 각각의 계층은 입력층(input layer), 숨은 층(hidden layer), 출력층(output layer)이다(Fig. 2).<sup>10,11)</sup> 학습을 하는 과정 중에서는 샘플 데이터가 입력 계층에 입력되고 예측된 출력층은 실제 결과와 비교된다. 인공 신경망의 예상 결과와 실제 결과는 오차가 발생할 수 있으며, 이 오차를 줄이기 위하여 결과에서 데이터 방향으로 반대 함수를 적용하여 오차량의 계수를 조정한다. 따라서 숨은 층은 특정 수학적 방법에 따라 가중치가 부여된 신경망(neuron)의 상호 연결을 수정하는 데 사용되는 역-전파(back-propagation)로서의 역할을 한다. 이를 통해 인공신경망이 연구중인 시스템의 모델을 최적으로 에뮬레이션(emulation)하여 결과적으로 훈련된 인공신경망을 사용하여 새로운 입력에 대한 결과를 예측하거나 추정할 수 있다.

## 2. 컨볼루션 신경망

일반적으로 이미지 처리는 인공신경망의 가장 큰 적용 분야 중 하나로, 컨볼루션 신경망은 척추 영상을 분석하는 기계학습 중 가장 유용한 방법으로 각광을 받고 있다. 인공신경망의 하위 개념인 컨볼루션 신경망은 영상 검사를 기계 학습을 통하여 분석할 수 있는 유용한 방법으로 컨볼루션 신경망의 설계는 고양이와 원숭이에서 수행된 실험을 기반으로 동물 시각 피질의 구조에 대한 연구에서부터 시작되었다.<sup>12)</sup> 1960년대에 Hubel과 Wiesel<sup>12)</sup>은

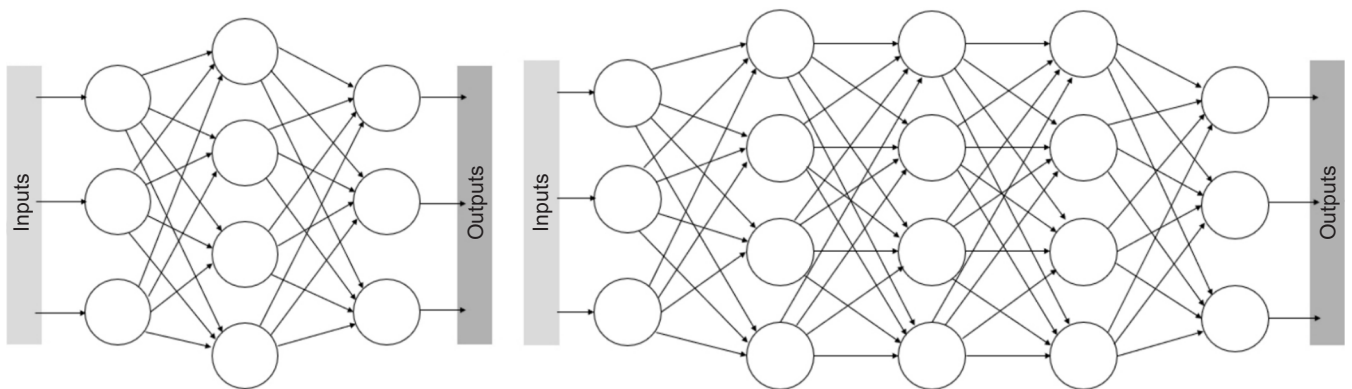


Figure 2. Schematic representation of an artificial neural network. An artificial neural network was composed of three layers of input, hidden, and output layer connecting each other. The difference between the expected and actual results of the artificial neural network serves as a backpropagation to modify the interconnections of the neural networks weighted according to a specific mathematical method. Cited from the article of Premalatha and Valan Arasu (J Appl Res Technol. 2016;14:206-14).<sup>11)</sup>

특정 뉴런 그룹이 시각 피질은 시야의 작은 영역에서만 자극을 받고 그 영역에서 특징과 정보를 추출하는 특성을 발표하였다. 특정 뉴런 그룹은 특정 부분의 이미지가 주는 정보에만 민감하고 기타 정보는 후 순위의 중요도를 보였으며, 다른 뉴런 그룹은 또 다른 정보에 민감하게 반응하는 특징을 보였다. 즉, 종합적인 시각적 인식은 서로 다른 뉴런 그룹에서 오는 정보를 결합하고 기존 정보에 대한 정보를 활용하여 도출하는 것으로, 컨볼루션 신경망은 이러한 신경 구조를 유사하게 모방한다.<sup>13)</sup>

컨볼루션 신경망에서 영상을 처리하여 인식하는 기본 특징은 영상의 명암, 색상 정보에 따라 각 층으로 나누고 픽셀은 특정 함수에 따라 숫자로 처리하는 것에서 시작한다. 척추 시상면 정렬의 분석을 위한 첫번째 단계는 대퇴 골두, 천추 및 척추의 상하부 종판을 인식하는 것으로, 컨볼루션 신경망에서 영상의 처리는 관심 영역의 픽셀은 강조하고 나머지 영역은 무시하는 알고리즘을 이용한다. 컨볼루션 층(convolutional layer)은 인식 가능한 작은 필터 단위로 일반적으로  $3 \times 3 \times 3$  또는  $5 \times 5 \times 3$ 의 크기를 가지며 영상의 종류에 따라 층의 단위가 달라진다. 이미지의 각 가능한 위치에 대해 숫자는 층 입력의 해당 값과 필터 가중치의 요소별 곱셈으로 계산되어 관심 영역에 접근하며 모든 숫자의 모음은 소위 활성화 지도를 구성한다(Fig. 3). 일반적인 컨볼루션 계층은 여러 층으로 구성되며 프로세스는 각 계층이 활성화되어 3차원 행렬을 구성한다. 분류 또는 회귀 작업을 통하여 인공 뉴런

의 미리 정의된 부분을 인위적으로 퇴거층(drop-out layer)을 취소하며 컨볼루션 신경망이 최적의 출력을 달성할 수 있도록 학습을 지속한다. 과도하게 함수의 차수를 높여 적합도를 올리려고 할 때 자료의 작은 변화 혹은 노이즈(noise)에 결과값이 영향을 받는 과적합(overfitting) 현상의 위험성을 줄이기 위하여 컨볼루션 계층 훈련은 필터 가중치의 최적 값을 찾는 것으로 구성되며 표준 인공신경망에 사용되는 것과 유사한 최적화 알고리즘을 통해 수행된다.<sup>14)</sup>

### 3. 인공 신경망의 척추 질환에 대한 적용

기계 학습 방법은 척추 질환에 대한 여러 연구에 적용되고 있다. 영상을 활용한 진단 및 치료 결과의 예측, 성인 척추 변형 수술의 범위 및 수술 전 계획을 결정하는데 인공 신경망이 활용된다. 특히 최근에는 컨볼루션 신경망을 활용한 척추 영상을 분석하는 다양한 시도들이 보고되어 있다.

#### 1) 진단

초기의 연구는 1997년 Bishop 등<sup>15)</sup>이 183명의 환자를 대상으로 다양한 범주의 척추 장애와 관련된 상체의 운동의 특징을 인공 신경망이 패턴을 구별할 수 있는지에 대한 연구였다. MLP 모델을 사용하였으며 3축 각도계로 동적 운동 데이터를 수집하여 161명의 환자에게 기계 학습을 시키고 22명의 환자에게 검증하

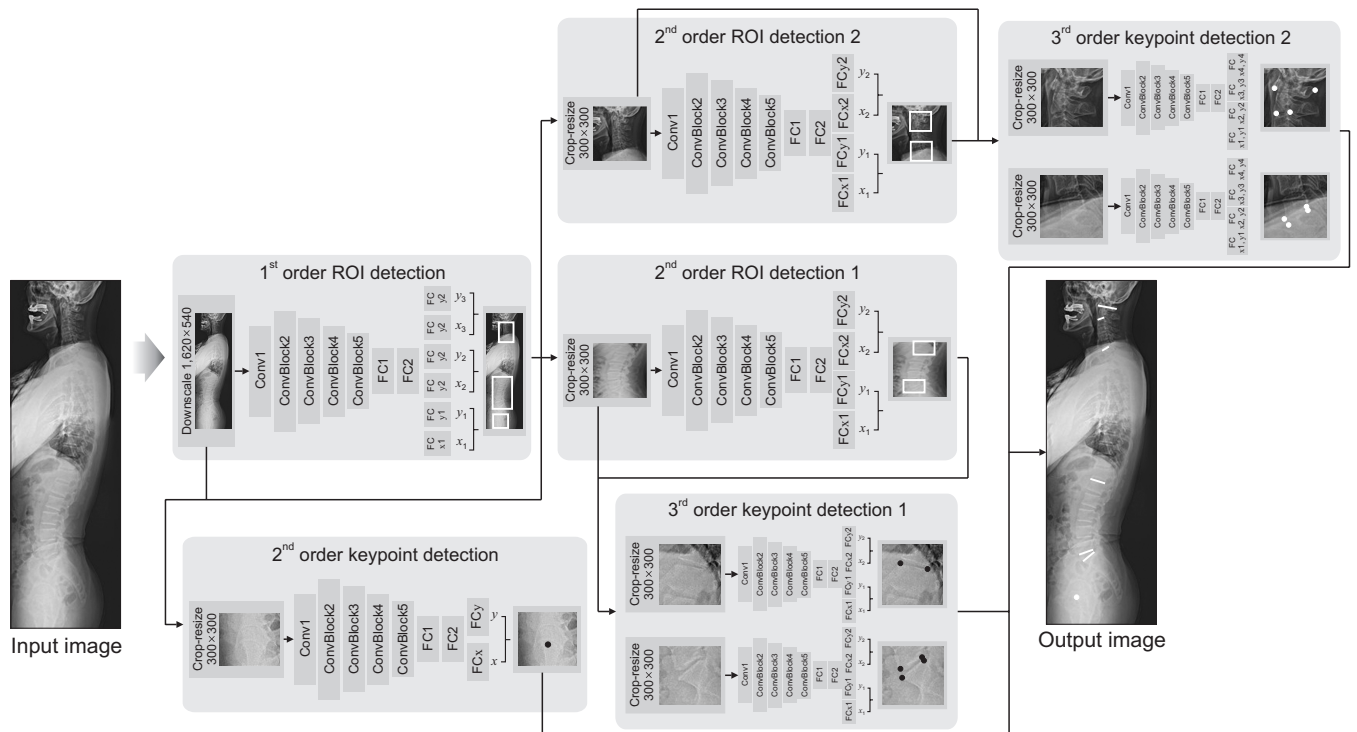


Figure 3. Example of the key process of the convolutional neural network to measure the sagittal spinal parameters automatically in a whole spine standing radiography (courtesy of professor Jonghun Yoon). ROI, region of interest.

Table 1. Brief Review of Research about Machine Learning Regarding the Diagnosis of Spinal Disorders

Author	Year	Model type	Training	Test	Objective	Results/conclusion
Bishop et al. <sup>15)</sup>	1997	MLP	161	22	To determine whether a neural network analysis system can be effective in distinguishing patterns.	The neural network classifier produced the best results with up to 85% accuracy on novel "validation" data.
Sari et al. <sup>17)</sup>	2012	ANN	169	169	To predict the intensity of low back pain using the ANN.	The suggested systems were found to be promising approaches in addition to existing unbiased approaches.
Kim et al. <sup>18)</sup>	2020	CNN	330	14	To develop web-based spine segmentation with deep learning using computed tomography scans.	The proposed web-based deep learning approach can be efficient and accurate for spine segmentation as a diagnostic method.
Kim et al. <sup>20)</sup>	2018	ANN	15,840 (70%)	6,789 (30%)	To train and validate machine learning models to identify risk factors for complications following posterior lumbar spine fusion.	Machine learning in logistic regression and ANNs was more accurate than benchmark American Society of Anesthesiology scores for identifying risk factors of developing complications following posterior lumbar spine fusion.
Lin <sup>22)</sup>	2008	MLP (ANN)	NA	NA	A multilayer feed-forward, backpropagation ANN was implemented to identify the classification patterns of the scoliosis spinal deformity.	The results in the study confirmed the existence of over-fitting or over-training in the selected training datasets in the validation learning session.
Zhang et al. <sup>23)</sup>	2017	CNN	235	105	To develop a computer-aided method that reduces the variability of Cobb angle measurement for scoliosis assessment.	For model radiographs, the intraclass correlation coefficients were greater than 0.98, and the mean absolute differences were less than 3°.
Lessmann et al. <sup>24)</sup>	2019	CNN	60	30	We propose an iterative instance segmentation approach that uses a fully convolutional neural network to segment and label vertebrae one after the other, independently of the number of visible vertebrae.	The anatomical identification had an accuracy of 93%, corresponding to a single case with mislabeled vertebrae. Vertebrae were classified as completely or incompletely visible with an accuracy of 97%.
Galbusera et al. <sup>25)</sup>	2019	CNN	443	50	To present an automated method for extracting anatomical parameters from biplanar radiographs of the spine	The standard errors of the estimated parameters ranged from 2.7° (for the pelvic tilt) to 11.5° (for the L1–L5 lordosis).
Pan et al. <sup>26)</sup>	2019	CNN	959	236	To automatically measure the Cobb angle and diagnose scoliosis on chest X-rays.	The computer-aided method achieved a high level of sensitivity (89.59%) and a relatively low level of specificity (70.37%) for diagnosing scoliosis on chest X-rays.
Weng et al. <sup>27)</sup>	2019	CNN (ResUNet)	900	90	To validate automatic SVA estimation on lateral spine radiography using the current deep-learning techniques	SVA measurement with a median absolute error of 1.183±0.166 mm within 0.2 seconds per image.
Yeh et al. <sup>28)</sup>	2021	CNN	2,210	400	Deep learning could be applied for automatic landmark detection and alignment analysis.	The proposed automatic alignment analysis system was able to localize spinal anatomic landmarks with high accuracy and to generate various radiographic parameters with favorable correlations with manual measurements.

MLP, multilayer perceptron; ANN, artificial neural network; CNN, convolutional neural network; NA, not assessable; SVA, sagittal vertical axis.

였다. 그들은 인공 신경망이 85%의 정확도로 요통을 분류할 수 있다고 하였다. 허부 요통을 분류하려는 연구는 이후에도 다양하게 시도되었고, 급성 혹은 만성 요통을 효과적으로 분류하였고, 예후를 예측하려는 연구가 보고되었다(Table 1).<sup>16-18)</sup>

자기공명영상 검사상 추간판의 퇴행성 변화에 대한 Modic 변화 및 Pfirrmann 등급을 분석한 기계 학습에 대한 연구는 Jamaludin 등<sup>19)</sup>의 연구가 대표적이다. 컨볼루션 신경망을 이용하여 1,806명의 환자에서 기계 학습시켰고, 203명의 환자에서 평가되었다. 이 인공 신경망은 추간판 감지 및 라벨링에서 95.6%의 정확도를 달성하여 객관성 및 분석 속도 측면에서 양호한 결과를 보고하였다.<sup>19)</sup> 동시기에 발표한 Kim 등<sup>20)</sup>, Han 등<sup>21)</sup>의 연구에서도 비슷한 수준의 정확성 및 일치도를 보고하였다.

Lin<sup>22)</sup>은 MLP 모델의 인공신경망을 활용하여 척추 측만증 모델의 패턴을 분류한 결과를 보고하였다. 17개의 입력층과 17개의 숨은 층의 뉴런 및 King의 분류에 따른 5개의 출력층으로 구성되었고, 척추 측만증 환자로부터 총 37개의 하위 패턴이 결정되었다. 이는 초기 예상한 5개의 출력층보다 높은 값으로 샘플 데이터의 부족으로 인한 것으로 판단된다고 하였다.<sup>22)</sup> 또한 Zhang 등<sup>23)</sup>은 척추 측만증 환자에서 콥스 각도를 기계학습으로 자동 측정하는 컨볼루션 신경망에 대한 분석을 보고하였다. 총 340명의 단순 방사선 검사 중 235명으로 기계 학습을 시키고 105명으로 시험을 하였을 때, 인공신경망과 척추 외과의의 수동 측정 차이는 5° 이상으로 오차가 크다고 보고하였다.<sup>23)</sup>

최근에는 컨볼루션 신경망을 통해 척추 정렬에 대한 자동 분석을 하였을 때 보다 향상된 결과를 보여주었다. Lessmann 등<sup>24)</sup>은 컨볼루션 신경망을 통한 척추 분절의 자동인식 학습 신경망을 보고하였고, 60명의 전산화 단층 촬영 검사로 학습시키고, 30명으로 평가하였을 시, 해부학적 특징의 식별률은 93%의 정확도를 보였고, 추체의 식별률은 97%의 정확도였다고 보고하였다. 한편 Galbusera 등<sup>25)</sup>은 추가로 미분할 수 있는 완전 컨볼루션 신경망을 바탕으로 흉추 후만각(T4-T12), 요추 전만각, 콥스 측만각, 골반 입사각, 천추 경사각 및 골반 경사각을 측정하는 기계 학습의 결과를 보고하였다. 443명에서 학습되었고, 50명에서 평가하였을 시 예측된 모든 매개 변수는 척추 외과의가 실측한 값과 강한 상관 관계가 있었고, 추정된 매개 변수의 표준 오차는 2.7° (골반 경사각)에서 11.5° (요추 전만각)까지 다양하다고 보고하였다.<sup>25)</sup> 최근에 보고된 다수의 연구들은 컨볼루션 신경망을 활용하여 콥스각(Cobb's angle) 및 척추 정렬의 지표의 측정에서 높은 관찰자 간 일치도(콥스각: 0.854, 시상 수직축: 0.946)를 보고하였다.<sup>26,27)</sup>

최근의 연구로 Yeh 등<sup>28)</sup>은 전척추 시상면 방사선 검사를 이용하여 45개의 해부학적 지표를 감지하고 주요 시상면 정렬 지표를 측정하는 컨볼루션 신경망 모델을 발표하였다. 2,200개의 평가 데이터셋을 활용하여 정확도를 높였으며 그 정확도는 해부학

적 지표의 탐지점의 개수와 관련이 있었다고 하였다. 즉, 2개의 해부학적 지표로 구성되는 경추와 요추의 지표의 정확도는 높은 반면, 4개의 지표로 구성되거나 흉추에서 흉곽 주위 구조물과의 구분이 어려운 흉추 및 골반 지표의 정확도는 상대적으로 낮았다. 컨볼루션 모델의 성능은 경추에서 가장 높았고 오차는 1.75-2.64 mm 범위였고, 그 다음으로 요추에서 1.76-2.63 mm 범위의 오차를 보였다. 흉추 부위는 더 큰 오차를 보였으며 오차 범위는 2.21-3.07 mm였고, 골반 지표의 오차는 가장 높았고 대퇴 골두의 중심의 측정에 대한 중심 오차가 3.39 mm였다고 보고하였다.<sup>28)</sup>

## 2) 결과 예측

Liszka-Hackzell과 Martin<sup>16)</sup>은 2002년도에 인공 신경망 기법을 이용하여 허부 요통의 불량한 예후가 예상되는 환자를 찾는 분석을 하였다. 15명의 급성 요통 환자와 25명의 만성 요통 환자가 등록되었으며 3주 동안 Mayo Clinic Electronic Symptom Diary를 이용하여 통증의 정도가 기록되었고, 분류 및 결과 데이터로 역전파 신경망을 훈련시켰다. 내원 각각 2, 6개월에 재평가하여 나쁜 예후를 보이는 환자를 인공 신경망을 통하여 예측하였다. 평가된 환자들은 일반 건강( $r=0.96$ ,  $p<0.01$ )과 정신 건강( $r=0.80$ ,  $p<0.01$ )에 대한 실제 값과 예측 값 사이에 좋은 상관 관계를 보여주어, 신경망 기술이 급성 및 만성 요통 환자를 분류하는데 효과적으로 적용될 수 있다고 보고하였다(Table 2).

Lafage 등<sup>29)</sup>은 성인 척추 변형에서 수술 전 후 단순 방사선 검사를 이용하여 상위 고정 척추를 선택하는 기계 학습법을 소개하였다. 143명의 전척추 시상면 단순 방사선 검사중에서 70%는 기계 학습에, 15%는 검증에 사용하였고, 나머지 15%에서 술 전 평가에 활용되었다. 인공 신경망에서 입력층은 8개, 숨은 층은 10개의 뉴런으로 구성되었고 흉추부의 상위 고정 척추를 예측하는 하나의 출력층으로 구성되었다. 훈련 후 결과는 테스트에서 81.0%의 정확도, 87.5%의 정밀도 및 87.5%의 재현율을 보여주어, 저자들은 인공신경망이 성인 척추 변형 환자의 교정술을 위한 2명의 척추 전문의의 결정을 성공적으로 모방이 가능하였다고 보고하였다.<sup>29)</sup>

또한 인공신경망은 요추 척추 협착증 및 요추 추간판 탈출증의 결과를 예측하고,<sup>30)</sup> 재발성 추간판 탈출증을 예측하는 데 우수한 결과를 보여주었다.<sup>31)</sup> 재발성 추간판 탈출증을 예측하는 Azimi 등<sup>31)</sup>의 연구에서는 총 402명의 환자가 각각 201명, 101명, 100명의 환자로 나누어 인공신경망의 훈련, 테스트 및 검증에 사용되었다. 추간판 탈출증의 재발률은 8.7%였고, 수술적 치료 후 수술 결과는 우수 31.1%, 좋음 44.3%, 보통 18.9%, 나쁨 5.7%로 분류되었다. 로지스틱 회귀 분석 모델과 비교할 때 인공신경망 모델은 우수한 결과를 예측하는 데 94.1%의 정확도로 예측 가능하였고, 곡선 아래 면적(area under curve, AUC)은 0.83으로 양

Table 2. A Brief Review of Research about Machine Learning Regarding the Prediction of Outcomes of Spinal Disorders

Author	Year	Model type	Training	Test	Objective	Results/conclusion
Liszka-Hackzell and Martin <sup>16)</sup>	2002	Backpropagation neural network	20	10	To present new techniques of patient assessment that may prospectively identify this minority of patients at risk of developing poor outcomes.	A good correlation showed between the true and predicted values for general health ( $r=0.96$ ) and mental health ( $r=0.80$ ).
Lafage et al. <sup>29)</sup>	2021	ANN	100	43	To train and validate an algorithm mimicking the decision-making of experienced surgeons regarding upper instrumented vertebra selection in surgical correction of thoracolumbar adult spinal deformity.	An artificial neural network successfully mimicked two lead surgeons' decision-making in selecting the upper-instrumented vertebra for adult spinal deformity correction.
Azimi et al. <sup>31)</sup>	2015	ANN	201	101	To develop an artificial neural network model to predict recurrent lumbar disk herniation.	The ANN model was associated with superior results: accuracy rate, 94.1%; Hosmer-Lemeshow statistic, 40.2%; and area under the curve, 0.83% of patients.
Cho et al. <sup>32)</sup>	2020	U-net	440 (70%)	189 (30%)	To develop a fully automated artificial intelligence and computer vision pipeline for assisted evaluation of lumbar lordosis.	The computer vision algorithm identified the L1 and S1 vertebrae on 84.1% of the test set with an average speed of 0.14 seconds/radiograph.
Azimi et al. <sup>33)</sup>	2017	ANN	174	86	To develop an ANN model that is designed to accurately select patients for surgery or non-surgical options in lumbar spinal stenosis.	The findings showed that an ANN model can predict the optimal treatment choice for lumbar spinal stenosis patients in a clinical setting and is superior to the logistic regression model.
Karhade et al. <sup>34)</sup>	2019	Machine learning algorithm	842 (80%)	210 (20%)	To develop machine learning algorithms for prediction of in-hospital and 90-day postdischarge mortality in spinal epidural abscess.	Machine learning algorithms show promise on internal validation for the prediction of 90-day mortality in a spinal epidural abscess.
Stopa et al. <sup>35)</sup>	2019	Machine learning algorithm	NA	144	To predict risk of nonhome discharge for patients undergoing surgery for lumbar disc disorders.	This external institutional validation of a previously developed machine-learning algorithm suggests a reliable method for identifying patients with lumbar disc disorder at risk for nonroutine discharge.
Zhang et al. <sup>36)</sup>	2019	Machine learning	58	22	To predict vertebral strength by using machine learning.	The maximum relative errors between vertebral strengths obtained from QCT and machine learning models were 7.733 and 6.958% (relative errors <7.740%).
Hopkins et al. <sup>37)</sup>	2020	DNN	3,034	1,012	To predict SSIs after posterior spinal fusions.	AUC=0.775, PPV=92.6%, NPV=98.5%.

ANN, artificial neural network; NA, not assessable; QCT, quantitative computed tomography; DNN, deep neural network; SSI, surgical site infection; AUC, the area under the curve; PPV, positive predictive value; NPV, negative predictive value.

호한 결과를 보여주었다.

Cho 등<sup>32)</sup>은 629명의 환자를 대상으로 U-net이라는 컨볼루션 신경망을 사용하여 요추부 단순 방사선 검사에서 해부학적 지표들 경추 전만각, 흉추 후만각, 골반 입사각과 요추 전만각

의 차이, 시상 수직축(sagittal vertical axis)을 포함한 주요 매개 변수를 측정하는 연구를 발표하였다. U-Net 컨볼루션 신경망은 ROC 곡선의 곡선 아래 영역 0.914, 관찰자 간 정확도 0.862를 보였고, 알고리즘은 평균 0.14초/영상 속도의 빠른 식별을

84.1%에서 보였다. 이 연구는 척추 외과와의 수동 측정 없이 킨 불루션 신경망으로 영상 자료를 식별한 것으로 임상적인 유용성을 제공하였다.<sup>32)</sup>

또한 인공신경망을 활용한 요추부 척추관 협착증 환자에 대한 수술적 치료 필요성을 예측하는 Azimi 등<sup>33)</sup>의 연구는 346명의 환자를 대상으로 시각통증척도와 Japanese Orthopaedic Association 척도, Oswestry 장애지수, 신경인성 파행결과 점수, 스위스 척추 협착증 점수를 인공신경망과 로지스틱 회귀모델을 비교하였다. 환자 정보는 훈련(n=174), 테스트(n=86) 및 검증(n=86) 데이터셋으로 구분되었고, 인공신경망 모델은 회귀 분석 모델보다 더 높은 97.8%의 정확도, Hosmer-Lemeshow 41.1%를 나타내어 더 잘 맞는 보정을 보였다(AUC=89.0%).<sup>33)</sup> 그 외에도 인공신경망 모델은 경막외 농양으로 수술적 치료를 받은 환자의 퇴원 후 90일 사망에 대한 예측,<sup>34)</sup> 요추 추간판 탈출증으로 수술적 치료를 받은 환자의 비계획적 퇴원의 예측,<sup>35)</sup> 고령의 척추 골강도의 평가 및 골다공증성 척추 골절 위험의 예측,<sup>36)</sup> 후방 척추 유합 후 수술 부위 감염을 예측하는 데 사용되었다.<sup>37)</sup>

#### 4. 인공 신경망의 제한점

인공 신경망을 이용한 의학 연구는 표본 크기 및 질에 의해 도출되는 결과가 영향을 받는다. 즉, 인공지능은 상식을 가지고 이성적인 판단을 하는 생명체의 지능이 아닌, 입력된 데이터를 토대로 하여 수학적 계산을 통하여 최적의 결과를 도출하려는 시도라는 것을 주지하여야 한다. 모든 인공지능에 의한 예측 모델의 성능은 훈련 데이터의 품질에 크게 좌우된다. 훈련 세트에 더 많은 데이터를 공급할수록 신경망은 특정 관심 결과를 더 잘 예측할 수 있는 특징을 보이며 보다 포괄적인 데이터셋을 사용하면 인공지능의 예측을 더욱 향상시킬 수 있다. 또한 인공 신경망은 알고리즘의 입력과 출력을 볼 수 있지만 모델의 내부 동작은 숨은 층으로 가려져 있기 때문에 인공 신경망이 예후를 추정할 수는 있으나 어떤 함수의 연산 과정을 거쳐 결과가 도출되었는지 세세히 알기는 힘들다. 이러한 이유로 인공신경망을 이용한 연구를 통하여 근본적인 병태 생리를 추정하는 것은 한계가 있다. 이러한 이유로 인공 신경망의 숨은 층의 연산은 블랙박스(black box)라고 일컫기도 한다. 초기의 인공지능의 중요성이 대두되었을 때, 의사의 역할을 인공지능이 대체할 것이라는 우려가 있었다. 그러나 이는 기우에 불과하며, 인공지능 기술의 의학에 대한 도입은 진료의 편의를 제공하며 단순 반복 작업을 줄여줄 수 있는 효율적인 도구라는 인식의 전환이 필요하다. 특히 컴퓨터의 그래픽 처리의 비약적인 발전으로 인하여 대규모 영상 데이터의 정량적인 분석을 가능하게 하여 진료의가 환자들에게 양질의 진료를 제공하는 데 도움을 주는 도구가 될 것으로 판단된다.

## 결론

기계 학습을 통한 인공지능을 의학 분야에 접목한 기술은 최근 급격한 성장을 보이고 있으며 이러한 경향은 향후 더 확대될 것으로 기대한다. 그러나 인공지능은 인간과 같은 상식과 이성적인 판단을 하는 명석한 두뇌가 아닌 입력된 데이터 셋을 토대로 빠르게 계산하는 컴퓨터 프로그래밍이다. 즉, 인공지능의 기술은 의료인을 대체할 수 있는 기술이 아닌 의료인에게 환자의 진단 및 치료에 대한 최신의 정보를 제공해 주는 유용한 도구인 것이다. 따라서 인공 신경망 기술을 의료로 발전시킬 수 있는 최신 기술로서 의료인이 주도적으로 도입하는 것이 중요하며, 특히 영상 자료가 환자의 진단 및 치료에 큰 영향을 미치는 정형외과 의사의 적극적인 관심과 참여가 중요할 것이다. 빠른 연산과 대규모 데이터를 신속히 처리할 수 있는 인공지능의 장점을 인지하여 의료정보화 시대에서 환자 진료와 치료에 인공지능을 적절히 활용하기를 기대한다.

## CONFLICTS OF INTEREST

The authors have nothing to disclose.

## ORCID

Sung Hoon Choi, <https://orcid.org/0000-0003-3328-2043>

Jonghun Yoon, <https://orcid.org/0000-0001-7547-3185>

Ye-Soo Park, <https://orcid.org/0000-0002-2801-2674>

## REFERENCES

1. Barrey C, Roussouly P, Le Huec JC, D'Acunzi G, Perrin G. Compensatory mechanisms contributing to keep the sagittal balance of the spine. *Eur Spine J.* 2013;22(Suppl 6):S834-41.
2. Alijani B, Rasoulian J. The sagittal balance of the cervical spine: radiographic analysis of interdependence between the occipitocervical and spinopelvic alignment. *Asian Spine J.* 2020;14:287-97.
3. Kim WJ, Shin HM, Song DG, et al. Comparison of clinical outcomes and complications of primary and revision surgery using a combined anterior and posterior approach in patients with adult spinal deformity and sagittal imbalance. *Clin Orthop Surg.* 2021;13:196-206.
4. Kim HJ, Yang JH, Chang DG, et al. Adult spinal deformity: current concepts and decision-making strategies for management. *Asian Spine J.* 2020;14:886-97.



5. Azimi P, Yazdani T, Benzel EC, et al. A review on the use of artificial intelligence in spinal diseases. *Asian Spine J.* 2020;14:543-71.
6. Azimi P, Benzel EC, Shahzadi S, Azhari S, Mohammadi HR. Use of artificial neural networks to predict surgical satisfaction in patients with lumbar spinal canal stenosis: clinical article. *J Neurosurg Spine.* 2014;20:300-5.
7. Samuel AL. Some studies in machine learning using the game of checkers. II—recent progress. *IBM J Res Dev.* 1967;11:601-17.
8. Galbusera F, Casaroli G, Bassani T. Artificial intelligence and machine learning in spine research. *JOR Spine.* 2019;2:e1044.
9. Hebb DO. The first stage of perception: growth of the assembly. *The organization of behavior: a neuropsychological theory.* New York: John Wiley & Sons; 1949. 60-78.
10. Parveen R, Nabi M, Memon FA, Zaman S, Ali M. A review and survey of artificial neural network in medical science. *J Adv Res Comput Appl.* 2016;3:7-16.
11. Premalatha N, Valan Arasu A. Prediction of solar radiation for solar systems by using ANN models with different back propagation algorithms. *J Appl Res Technol.* 2016;14:206-14.
12. Hubel DH, Wiesel TN. Receptive fields and functional architecture of monkey striate cortex. *J Physiol.* 1968;195:215-43.
13. Denker JS, Gardner WR, Graf HP, et al. Neural network recognizer for hand-written zip code digits. Paper presented at: 2nd IEEE Conference on Neural Information Processing Systems; 1988; Denver, USA.
14. LeCun Y, Boser B, Denker JS, et al. Backpropagation applied to handwritten zip code recognition. *Neural Comput.* 1989;1:541-51.
15. Bishop JB, Szpalski M, Ananthraman SK, McIntyre DR, Pope MH. Classification of low back pain from dynamic motion characteristics using an artificial neural network. *Spine (Phila Pa 1976).* 1997;22:2991-8.
16. Liszka-Hackzell JJ, Martin DP. Categorization and analysis of pain and activity in patients with low back pain using a neural network technique. *J Med Syst.* 2002;26:337-47.
17. Sari M, Gulbandilar E, Cimbiz A. Prediction of low back pain with two expert systems. *J Med Syst.* 2012;36:1523-7.
18. Kim YJ, Ganbold B, Kim KG. Web-based spine segmentation using deep learning in computed tomography images. *Healthc Inform Res.* 2020;26:61-7.
19. Jamaludin A, Lootus M, Kadir T, et al. ISSLS PRIZE IN BIO-ENGINEERING SCIENCE 2017: automation of reading of radiological features from magnetic resonance images (MRIs) of the lumbar spine without human intervention is comparable with an expert radiologist. *Eur Spine J.* 2017;26:1374-83.
20. Kim JS, Merrill RK, Arvind V, et al. Examining the ability of artificial neural networks machine learning models to accurately predict complications following posterior lumbar spine fusion. *Spine (Phila Pa 1976).* 2018;43:853-60.
21. Han Z, Wei B, Mercado A, Leung S, Li S. Spine-GAN: semantic segmentation of multiple spinal structures. *Med Image Anal.* 2018;50:23-35.
22. Lin H. Identification of spinal deformity classification with total curvature analysis and artificial neural network. *IEEE Trans Biomed Eng.* 2008;55:376-82.
23. Zhang J, Li H, Lv L, Zhang Y. Computer-aided Cobb measurement based on automatic detection of vertebral slopes using deep neural network. *Int J Biomed Imaging.* 2017;2017:9083916.
24. Lessmann N, van Ginneken B, de Jong PA, Išgum I. Iterative fully convolutional neural networks for automatic vertebra segmentation and identification. *Med Image Anal.* 2019;53:142-55.
25. Galbusera F, Niemeyer F, Wilke HJ, et al. Fully automated radiological analysis of spinal disorders and deformities: a deep learning approach. *Eur Spine J.* 2019;28:951-60.
26. Pan Y, Chen Q, Chen T, et al. Evaluation of a computer-aided method for measuring the Cobb angle on chest X-rays. *Eur Spine J.* 2019;28:3035-43.
27. Weng CH, Wang CL, Huang YJ, et al. Artificial intelligence for automatic measurement of sagittal vertical axis using ResUNet framework. *J Clin Med.* 2019;8:1826.
28. Yeh YC, Weng CH, Huang YJ, et al. Deep learning approach for automatic landmark detection and alignment analysis in whole-spine lateral radiographs. *Sci Rep.* 2021;11:7618.
29. Lafage R, Ang B, Alshabab BS, et al. Predictive model for selection of upper treated vertebra using a machine learning approach. *World Neurosurg.* 2021;146:e225-32.
30. Azimi P, Benzel EC, Shahzadi S, Azhari S, Mohammadi HR. The prediction of successful surgery outcome in lumbar disc herniation based on artificial neural networks. *J Neurosurg Sci.* 2016;60:173-7.
31. Azimi P, Mohammadi HR, Benzel EC, Shahzadi S, Azhari S. Use of artificial neural networks to predict recurrent lumbar disk herniation. *J Spinal Disord Tech.* 2015;28:E161-5.
32. Cho BH, Kaji D, Cheung ZB, et al. Automated measurement

- of lumbar lordosis on radiographs using machine learning and computer vision. *Global Spine J.* 2020;10:611-8.
33. Azimi P, Mohammadi HR, Benzel EC, Shahzadi S, Azhari S. Use of artificial neural networks to decision making in patients with lumbar spinal canal stenosis. *J Neurosurg Sci.* 2017;61:603-11.
34. Karhade AV, Shah AA, Bono CM, et al. Development of machine learning algorithms for prediction of mortality in spinal epidural abscess. *Spine J.* 2019;19:1950-9.
35. Stopa BM, Robertson FC, Karhade AV, et al. Predicting non-routine discharge after elective spine surgery: external validation of machine learning algorithms. *J Neurosurg Spine.* 2019;31:742-7.
36. Zhang M, Gong H, Zhang K, Zhang M. Prediction of lumbar vertebral strength of elderly men based on quantitative computed tomography images using machine learning. *Osteoporos Int.* 2019;30:2271-82.
37. Hopkins BS, Mazmudar A, Driscoll C, et al. Using artificial intelligence (AI) to predict postoperative surgical site infection: a retrospective cohort of 4046 posterior spinal fusions. *Clin Neurol Neurosurg.* 2020;192:105718.

# 척추 질환의 진단 및 결과 예측 분야에 대한 인공지능 학습법의 적용

최성훈 · 윤종현\* · 박예수<sup>†</sup>

한양대학교 의과대학 한양대학교병원 정형외과학교실, \*한양대학교 공과대학 ERICA 기계공학부,  
†한양대학교 의과대학 한양대학교구리병원 정형외과학교실

인공 지능은 기계 학습 및 인공 신경망을 포함하는 개념이며, 컴퓨터의 빠른 연산을 바탕으로 대규모 데이터의 신속한 분석이 가능한 특징을 보인다. 인공 지능 기술을 도입한 특정 질환의 진단 및 예후를 추정하는 다양한 의학 연구가 보고되었다. 최근에는 영상 자료의 인식 및 분석에 유용한 컨볼루션 신경망(convolutional neural network) 학습법을 활용한 많은 의학 관련 연구가 보고되어 있다. 척추 질환을 포함한 정형외과의 다양한 질환들은 단순 방사선 검사, 전산화 단층 촬영 검사 및 자기 공명 영상 검사 등 영상 자료가 환자의 진단 및 치료를 결정하는 데 필수적이다. 따라서 타 의학 분야에 비해 정형외과 질환의 진단 및 치료에 인공 지능 학습법이 적용될 수 있는 많은 영역이 있다. 의료 정보화 시대에 정형외과 의사가 주도적으로 인공 지능 기술을 도입하는 것은 환자의 진단 및 치료에 긍정적인 요인이 되어 의학 발전에 기여하는 바가 클 것으로 판단된다. 본 종설에서는 인공 지능의 개략적인 내용을 정리하고 척추 질환의 진단 및 결과 예측에 기계 학습법을 활용한 연구들을 소개한다.

**색인단어:** 기계학습, 인공 신경망, 컨볼루션 신경망, 척추, 진단

접수일 2021년 7월 12일 수정일 2021년 9월 5일 게재확정일 2021년 10월 13일

†책임저자 박예수

11923, 구리시 경춘로 153, 한양대학교 구리병원 정형외과

TEL 031-560-2187, FAX 031-557-8781, E-mail [hyparkys@hanyang.ac.kr](mailto:hyparkys@hanyang.ac.kr), ORCID <https://orcid.org/0000-0002-2801-2674>